FEEDBACK DATA RATES FOR NONLINEAR SYSTEMS

Girish N. Nair, Robin J. Evans, Iven M.Y. Mareels, William Moran

Department of Electrical and Electronic Engineering
University of Melbourne, VIC 3010, Australia
Email: {g.nair, r.evans, i.mareels, bmor}@ee.mu.oz.au

Keywords: stabilizability, communication channels, entropy

Abstract

This paper poses a simple question: what is the lowest rate, in bits per unit time, at which feedback information can be transmitted in order to stabilise a given dynamical system? Expressions for this fundamental quantity have recently been derived for linear systems, with and without noise. In this work, the case of deterministic, fully observed, continuously differentiable dynamical systems is investigated, under the additional assumptions of controllability to the desired set-point and bounded initial states. By the use of volume-partitioning arguments and local Jordan forms, the infimum feedback data rate is shown to be the base-2 logarithm of the magnitude of the determinant of the open-loop Jacobian on the local unstable subspace, evaluated at the set-point. Connections to the concept of local topological feedback entropy are briefly discussed.

1 Introduction

In many developing application areas such as micro-electromechanical systems and decentralized tracking, the resources available for communication between sensors and controllers can be severely limited, due to size or cost. This impinges directly on the feedback control performance that can be achieved, since it implies that the data received by various components is either out-of-date or poor in resolution, if not both. In these situations the communications and control issues are intimately related and the analysis of one aspect cannot proceed without consideration of the other.

The focus in this paper is on communication constraints that take the form of limited data rates, in bits per unit time. In particular, the aim is to investigate the stabilisability of a given dynamical system when feedback information is received over a noiseless, digital channel with a finite data rate. A wealth of results are available for linear, time-invariant systems without disturbances, starting with the seminal paper [3] and continuing with [17, 1, 4, 2, 13, 8, 5]. Particularly relevant are [9, 15, 11, 1, 6], in which necessary *and* sufficient data rate bounds for the stabilisability of noiseless linear systems were derived, and [10], in which a tight bound for stochastic linear systems was derived. Despite different formulations and assumptions, the bounds obtained were generally all equal to the sum of the logarithms of the magnitudes of the unstable, open-loop eigenvalues.

This line of inquiry is extended here towards a general class of nonlinear systems. Assuming that the dynamical map is continuously differentiable and has a fixed point, the objective is to find the infimum data rate above which there exists a coding and control law which uniformly exponentially stabilises the system. In the next section, the problem is formulated precisely and the main result, Theorem 1 stated. Its connections to the open-loop notions of *Kolmogorov-Sinai* and *topological* entropy are also briefly discussed. The rest of the paper is the proof of the theorem. In section 3, a volume-partitioning argument is used to establish the necessity of the data rate bound specified. Its sufficiency is then confirmed in the penultimate section by explicitly constructing a coding and control scheme and establishing uniform exponential stability.

2 Formulation and Statement of Main Result

First, certain conventions need to be defined. Sequences $\{a_j\}_{j=0}^k$ are denoted $\tilde{a}_k$ and $\|\cdot\|$ represents either the Euclidean norm on a vector space or the matrix norm induced by it. Lebesgue measure is denoted λ, vectors are written in bold-face type and matrices in bold-face upper-case. The $n \times n$ identity matrix is denoted by $\mathbf{I}_n$, the $m \times n$ 0 matrix by $\mathbf{0}_{m\times n}$, the $m \times n$ matrix with 1's on the left upper-corner diagonal and zeros elsewhere by $\mathbf{I}_{m\times n}$ and the spectrum of a matrix is represented as $\sigma(\cdot)$, with multiple eigenvalues permitted. As usual, the real numbers are written $\mathbb{R}$, complex numbers $\mathbb{C}$, positive integers $\mathbb{N}$ and non-negative integers $\mathbb{Z}_+$.

Consider the fully observed, nonlinear, time-invariant system

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k), \quad \forall k \in \mathbb{Z}_+, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state and $\mathbf{u}_k \in \mathbb{R}^m$ the control vector. It is assumed that

A1 the dynamical map $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is differentiable once with continuous 1st order partial derivatives,

A2 there exist a fixed point $\mathbf{x}_*$ and constant input $\mathbf{u}_*$ such that $\mathbf{x}_* = \mathbf{f}(\mathbf{x}_*, \mathbf{u}_*)$,

A3 $\mathbf{f}$ is controllable to $\mathbf{z}_*$ in the sense that given any $l, \varepsilon > 0$, $\exists N \in \mathbb{N}, U > 0$ s.t. $\forall \|\mathbf{x}_0 - \mathbf{x}_*\| \leq l$, there exists a control sequence $\{\mathbf{u}_k\}_{k=0}^{N-1}$ with $\|\mathbf{u}_k - \mathbf{u}_*\| \leq U$ that ensures $\|\mathbf{x}_N - \mathbf{x}_*\| \leq \varepsilon$,

A4 the pair $(\mathbf{A}, \mathbf{B})$ is controllable, where $\mathbf{A}$ and $\mathbf{B}$ are the Jacobians of $\mathbf{f}$ w.r.t. state and control respectively at $(\mathbf{x}_*, \mathbf{u}_*)$.

Suppose that the sensor measuring the states is connected to a distant controller by a noiseless digital channel that can carry only one discrete-valued symbol s_k per sampling interval, selected from a coding alphabet $\mathcal{S}_k$ of time-varying size μ_k. The transmission data rate R may then be defined as the asymptotic average bit rate

$$R \triangleq \liminf_{k\to\infty} k^{-1}\sum_{j=0}^{k-1}\log_2 \mu_j. \tag{2}$$

This is a more general definition of channel data rate than used in the previous work on linear systems [11] and the motivation for its use will become clear in section 4.

As the objective here is not to address computational limitations, each symbol transmitted by the coder is permitted to depend on all past and present measurements and past symbols, i.e.

$$s_k = \gamma_k(\tilde{\mathbf{x}}_k, \tilde{s}_{k-1}), \quad \forall k \in \mathbb{Z}_+, \tag{3}$$

where $\gamma_k : \mathbb{R}^{n\times(k+1)} \times \tilde{\mathcal{S}}_{k-1} \to \mathcal{S}_k$ is the coder mapping at time k. It is assumed that for any sequence $\tilde{c}_{k-1} \in \tilde{\mathcal{S}}_{k-1}$, the coding partitions $\{\gamma_k^{-1}(c_k, \tilde{c}_{k-1}) \subset \mathbb{R}^{n\times(k+1)}\}_{c_k\in\mathcal{S}_k}$ are measurable. Neglecting transmission errors, assume that each symbol takes one sampling interval to be completely transmitted. Hence at time k the controller has $s_0, \ldots, s_{k-1}$ available and generates

$$\mathbf{u}_k = \delta_k(\tilde{s}_{k-1}), \quad \forall k \in \mathbb{Z}_+, \tag{4}$$

where $\delta_k : \tilde{\mathcal{S}}_{k-1} \to \mathbb{R}^m$ is the controller function at time k.

Define the *coder-controller* as the triplet of alphabet, coder and controller sequences $(\tilde{\mathcal{S}}_\infty, \tilde{\gamma}_\infty, \tilde{\delta}_\infty)$. Given an asymptotic average data rate $R > 0$, the primary objective here is to investigate whether there exists one that uniformly exponentially stabilises the plant (1) over initial states within a ball of radius l_0, i.e.

$$\varrho^{-k} \sup_{\|\mathbf{x}_0\|\le l_0} \{\|\mathbf{x}_k - \mathbf{x}_*\|, \|\mathbf{u}_k - \mathbf{u}_*\|\} \to 0 \text{ as } k \to \infty \tag{5}$$

for some $\varrho \in (0,1)$, where $\mathbf{x}_*, \mathbf{u}_*$ are the fixed point and constant input defined in assumption A2. The weaker notion of uniform asymptotic stability, corresponding to $\varrho = 1$, will also be explored.

For finite-dimensional, linear systems, it is known that there is a critical data rate which determines whether closed-loop stability is possible or not [15, 11]. It may therefore be expected that a critical rate will also exist for the case of an nonlinear system. The main result of this paper is now stated:

Theorem 1 *Let assumptions A1–A4 hold for the plant (1). Then any coder-controller (3)-(4) which stabilises the plant in the uniform exponential sense (5) must have asymptotic average feedback data rate R (in bits per interval) (2) strictly satisfying*

$$R > \sum_{\eta\in\sigma(\mathbf{A}):\, |\eta|\ge 1} \log_2|\eta|, \tag{6}$$

where $\mathbf{A}$ is the Jacobian of the dynamical map $\mathbf{f}$ with respect to state, evaluated at the set-point, and $\sigma(\mathbf{A})$ is the spectrum of $\mathbf{A}$.

Furthermore this bound is tight, i.e. for any number R_0 satisfying it there exists a uniformly exponentially stabilising coder-controller that has a smaller data rate.

For the weaker notion of uniform asymptotic stability this lower bound is still necessary, but equality may be possible.

As no assumptions apart from causality and measurability have been placed on the coding and control scheme, this in a very general sense defines the infimum rate at which information can circulate in a stable, deterministic feedback loop. Note that is a function only of the local, open-loop dynamics at the desired set-point and agrees with the results of [15, 11] for linear systems without disturbances.

As in the linear case, insight into the meaning of (6) can be obtained by rewriting it as $2^R > \prod_{|\eta|\ge 1}|\eta|$. The RHS is the factor by which a volume in the locally unstable subspace increases at each time step in open-loop, while the left-hand side is the average number of coding regions into which this volume can be partitioned. Hence the system is stabilisable if and only if the increase in unstable uncertainty volume can be counteracted by the decrease due to coding.

The RHS of (6) is strikingly similar to expressions for the *Kolmogorov-Sinai* and *topological* entropy rates of linear maps; see e.g. [16]. The crucial difference with the result above is that these notions are defined for open-loop systems, whereas the infimum stabilising data rate is a closed-loop concept. Nonetheless, it is possible to rigorously define a *topological feedback entropy (TFE)* for the plant, describing the rate at which the plant generates information in a stable feedback loop [12]. By taking appropriate limits a *local* TFE at a fixed point can then be defined and shown to coincide with the RHS of (6). From this viewpoint, exponential stability is possible if and only if the data rate of the channel exceeds the local TFE at the desired set-point, an interpretation that parallels Shannon's source coding theorem in digital communications [14]. However, the remainder of this paper is devoted to proving Theorem 1 without referring to the notion of TFE.

3 Necessity

As mentioned above, the intuition behind the necessity of (6) is that the open-loop growth in unstable subspace uncertainty volume near the set-point must be counteracted by the reduction in volume due to the coding partitions. A similar idea was employed in the linear case [15, 11]. However, the nonlinearity of the plant necessitates rather different technical tools.

Suppose that uniform exponential stability has been achieved by some coder-controller $(\tilde{\mathcal{S}}_\infty, \tilde{\gamma}_\infty, \tilde{\delta}_\infty)$. Recall that $\mathbf{A}$ is the Jacobian of the dynamical map w.r.t. the state at the set-point,

$$\mathbf{A} \triangleq \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{(\mathbf{x}_*,\mathbf{u}_*)} \in \mathbb{R}^{n\times n}, \tag{7}$$

and let $\mathbf{T} \in \mathbb{R}^{n\times n}$ be an orthonormal real similarity transform such that

$$\mathbf{J} \triangleq \mathbf{T}\mathbf{A}\mathbf{T}' \in \mathbb{R}^{n\times n} \tag{8}$$

is a *real Jordan form*; see e.g. [7] for details. Briefly, $\mathbf{J}$ has a block-diagonal structure with each block possessing either one real or two complex conjugate eigenvalues, not counting repeats. In terms of system dynamics, $\mathbf{z}_k \triangleq \mathbf{T}\mathbf{x}_k$ can then be interpreted as a vector of modes with decoupled open-loop dynamics near the set-point.

Define $\mathbf{z}_k^{\mathrm{u}} \in \mathbb{R}^d$ to be the vector of those modes governed by eigenvalues of $\mathbf{J}$ not less than 1 in magnitude. Assuming without loss of generality that the blocks of $\mathbf{J}$ are ordered according to descending eigenvalue magnitudes,

$$\mathbf{z}_k^{\mathrm{u}} = \mathbf{I}_{d\times n}\mathbf{x}_k, \quad \text{with } \mathbf{z}_*^{\mathrm{u}} \triangleq \mathbf{I}_{d\times n}\mathbf{x}_*. \tag{9}$$

It then trivially follows that $\|\mathbf{x}_k - \mathbf{x}_*\| = \|\mathbf{z}_k - \mathbf{z}_*\| \geq \|\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}\|$, so that $\mathbf{z}_k^{\mathrm{u}} \to \mathbf{z}_*^{\mathrm{u}}$ exponentially in k and uniformly over $\mathbf{x}_0$.

Next, $\forall \tilde{c}_{k-2} \in \tilde{\mathcal{S}}_{k-2}$ define the *locally unstable uncertainty set*

$$I_k(\tilde{c}_{k-2}) \triangleq \left\{\mathbf{z}^{\mathrm{u}} \in \mathbb{R}^d |\, \|\mathbf{x}_0\| \leq l_0, \tilde{s}_{k-2} = \tilde{c}_{k-2}, \mathbf{z}^{\mathrm{u}} = \mathbf{z}_k^{\mathrm{u}}\right\}, \tag{10}$$

i.e. the set of all possible points that $\mathbf{z}_k^{\mathrm{u}}$ can take given the symbol sequence $\tilde{s}_{k-2} = \tilde{c}_{k-2}$. As the dynamical map $\mathbf{f}$ is continuous and the coding partitions measurable, these uncertainty sets are also measurable and so a *worst-case locally unstable uncertainty volume*

$$v_k \triangleq \max_{\tilde{c}_{k-2}} \lambda\{I_k(\tilde{c}_{k-2})\} \tag{11}$$

can be defined. Now, if r denotes the supremum distance of points in a measurable set $H \subset \mathbb{R}^d$ from the distance, then H is obviously wholly contained in the ball of radius r centred at the origin. Hence

$$\lambda\{H\} \leq \beta r^d = \beta \sup_{\mathbf{x}\in H} \|\mathbf{x}\|^d, \ \forall \text{ measurable } H \subset \mathbb{R}^d,$$

where β is the d-dimensional sphere constant. Thus

$$\begin{aligned}
l_k &\triangleq \sup_{\|\mathbf{x}_0\|\leq l_0} \{\|\mathbf{z}_k - \mathbf{z}_*\|, \|\mathbf{u}_k - \mathbf{u}_*\|\} \geq \sup_{\|\mathbf{x}_0\|\leq l_0} \|\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}\|, \\
&= \max_{\tilde{c}_{k-2}} \sup_{\|\mathbf{x}_0\|\leq l_0,\, \tilde{s}_{k-2}=\tilde{c}_{k-2}} \|\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}\|, && (12)\\
&\geq \max_{\tilde{c}_{k-2}} \beta^{-1/d} \lambda\{\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}} |\, \|\mathbf{x}_0\| \leq l_0,\ \tilde{s}_{k-2} = \tilde{c}_{k-2}\}^{1/d}, \\
&= \beta^{-1/d} \max_{\tilde{c}_{k-2}} \lambda\{\mathbf{z}_k^{\mathrm{u}} |\, \|\mathbf{x}_0\| \leq l_0,\ \tilde{s}_{k-2} = \tilde{c}_{k-2}\}^{1/d}, && (13)\\
&= \beta^{-1/d} \max_{\tilde{c}_{k-2}} \lambda\{I_k(\tilde{c}_{k-2})\} \equiv \beta^{-1/d} v_k^{1/d}, && (14)
\end{aligned}$$

i.e., $v_k \to 0$ exponentially as well. The equality (13) is a consequence of the invariance of Lebesgue measure to constant translations, while the equality in (12) follows from the fact that, with the coder-controller fixed, the same $\mathbf{x}_0$ cannot yield two different symbol sequences, i.e. the regions $\{\mathbf{x}_0 \in \mathbb{R}^n : \|\mathbf{x}_0\| \leq l_0,\, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}_{\tilde{c}_{k-2}\in\tilde{\mathcal{S}}_{k-2}}$ must be disjoint and exhaustive.

A recursive lower bound for the worst-case volume v_k will now be derived. Observe that

$$\begin{aligned}
v_{k+1} &\triangleq \max_{\tilde{c}_{k-1}} \lambda\{\mathbf{z}_{k+1}^{\mathrm{u}} |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\}, \\
&= \max_{\tilde{c}_{k-1}} \lambda\left\{\mathbf{g}(\mathbf{z}_k^{\mathrm{u}}, \mathbf{z}_k^{\mathrm{s}}, \mathbf{u}_k) |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\right\}, && (15)
\end{aligned}$$

where for convenience, the locally stable components of $\mathbf{T}\mathbf{x}_k$ are denoted $\mathbf{z}_k^{\mathrm{s}} \in \mathbb{R}^{n-d}$ and $\mathbf{g}(\mathbf{z}^{\mathrm{u}}, \mathbf{z}^{\mathrm{s}}, \mathbf{u}) \triangleq \mathbf{I}_{d\times n}\mathbf{T}\mathbf{f}(\mathbf{x}, \mathbf{u})$. The next step is to replace the nonlinear function $\mathbf{g}$ with its local linearisation. As $\mathbf{f}$ has continuous first order derivatives, $\forall \|\mathbf{z} - \mathbf{z}_*\|, \|\mathbf{u} - \mathbf{u}_*\| \leq l$

$$\mathbf{g}(\mathbf{z}^{\mathrm{u}}, \mathbf{z}^{\mathrm{s}}, \mathbf{u}) = \mathbf{z}_*^{\mathrm{u}} + \mathbf{J}^{\mathrm{u}}(\mathbf{z}^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}) + \mathbf{I}_{d\times n}\mathbf{T}\mathbf{B}(\mathbf{u} - \mathbf{u}_*) + \mathrm{o}(l)$$

uniformly over $\mathbf{z}, \mathbf{u}$ as $l \to 0$, where $\mathbf{J}^{\mathrm{u}} \triangleq \mathbf{I}_{d\times n}\mathbf{T}\mathbf{J}\mathbf{T}'\mathbf{I}_{n\times d} \in \mathbb{R}^{d\times d}$, the Jordan form governing the locally unstable subspace. From this it can established that $\exists\varepsilon(l) \to 0$ s.t. for any measurable $H \subset \{(\mathbf{z}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^m |\, \|\mathbf{z} - \mathbf{z}_*\|, \|\mathbf{u} - \mathbf{u}_*\| \leq l\}$,

$$\begin{aligned}
&\lambda\left\{\mathbf{g}(\mathbf{z}^{\mathrm{u}}, \mathbf{z}^{\mathrm{s}}, \mathbf{u}) |\, (\mathbf{z}, \mathbf{u}) \in H\right\} \geq \\
&\qquad [1 - \varepsilon(l)]\lambda\left\{\mathbf{z}_*^{\mathrm{u}} + \mathbf{J}^{\mathrm{u}}(\mathbf{z}^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}) + \mathbf{I}_{d\times n}\mathbf{T}\mathbf{B}\mathbf{u} |\, (\mathbf{z}, \mathbf{u}) \in H\right\}
\end{aligned}$$

Substituting this into (15) with $l = l_k$, $\mathbf{z}_k^{\mathrm{u}} = \mathbf{z}^{\mathrm{u}}$, $\mathbf{z}_k^{\mathrm{s}} = \mathbf{z}^{\mathrm{s}}$, $\mathbf{u}_k = \mathbf{u}$ and $\{(\mathbf{z}_k, \mathbf{u}_k) |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\} = H$, and writing $\varepsilon(l_k) \equiv \varepsilon_k$,

$$\begin{aligned}
&v_{k+1} \\
&\quad\geq (1 - \varepsilon_k) \max_{\tilde{c}_{k-1}} \lambda\left\{\mathbf{z}_*^{\mathrm{u}} + \mathbf{J}^{\mathrm{u}}(\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}})\right. \\
&\qquad\qquad \left. + \mathbf{I}_{d\times n}\mathbf{T}\mathbf{B}\delta_k(\tilde{c}_{k-1}) |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\right\}, \\
&\quad= (1 - \varepsilon_k) \max_{\tilde{c}_{k-1}} \lambda\left\{\mathbf{J}^{\mathrm{u}}(\mathbf{z}_k^{\mathrm{u}} - \mathbf{z}_*^{\mathrm{u}}) |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\right\}, && (16)\\
&\quad= (1 - \varepsilon_k) \max_{\tilde{c}_{k-1}} |\det \mathbf{J}^{\mathrm{u}}| \lambda\left\{\mathbf{z}_k^{\mathrm{u}} |\, \tilde{s}_{k-1} = \tilde{c}_{k-1}\right\}, && (17)\\
&\quad\equiv (1 - \varepsilon_k) |\det \mathbf{J}^{\mathrm{u}}| \max_{\tilde{c}_{k-2}} \left\{ \max_{c_{k-1}} \lambda\left\{\mathbf{z}_k^{\mathrm{u}} |\, s_{k-1} = c_{k-1},\right.\right. \\
&\qquad\qquad \left.\left. \tilde{s}_{k-2} = \tilde{c}_{k-2}\right\}\right\}, && (18)
\end{aligned}$$

where (16) follows from the translation-invariance of Lebesgue measure and (17) describes the effect of an invertible linear transformation on volume.

The trivial decomposition (18) leads to an observation that is the heart of the necessity argument developed here. The uncertainty sets $\{\mathbf{z}_k^{\mathrm{u}} |\, s_{k-1} = c_{k-1}, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}$ are not necessarily disjoint as the single symbol c_{k-1} runs over its possible values. However, since s_{k-1} is a well-defined function of the initial state and previous symbols, their union must cover the set $\{\mathbf{z}_k^{\mathrm{u}} | \tilde{s}_{k-2} = \tilde{c}_{k-2}\}$, i.e.

$$\begin{aligned}
&\lambda\left(\{\mathbf{z}_k^{\mathrm{u}} |\, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}\right) \\
&\quad= \lambda\left(\bigcup_{c_{k-1}=0}^{\mu_{k-1}-1} \{\mathbf{z}_k^{\mathrm{u}} |\, s_{k-1} = c_{k-1}, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}\right), \\
&\quad\leq \sum_{c_{k-1}=0}^{\mu_{k-1}-1} \lambda\{\mathbf{z}_k^{\mathrm{u}} |\, s_{k-1} = c_{k-1}, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}, \\
&\quad\leq \mu_{k-1} \max_{c_{k-1}} \lambda\{\mathbf{z}_k^{\mathrm{u}} |\, s_{k-1} = c_{k-1}, \tilde{s}_{k-2} = \tilde{c}_{k-2}\}. && (19)
\end{aligned}$$

Substituting this into (18),

$$\begin{aligned} v_{k+1} &\geq (1-\varepsilon_k)|\det \mathbf{J}^{\mathrm{u}}| \max_{\tilde{c}_{k-2}} \mu_{k-1}^{-1} \lambda \left\{ \mathbf{z}_k^{\mathrm{u}} | \, \tilde{s}_{k-2} = \tilde{c}_{k-2} \right\}, \\ &= \frac{(1-\varepsilon_k)|\det \mathbf{J}^{\mathrm{u}}|}{\mu_{k-1}} v_k \geq v_1 \prod_{j=1}^{k} \frac{(1-\varepsilon_j)|\det \mathbf{J}^{\mathrm{u}}|}{\mu_{j-1}}, \end{aligned}$$

by repeating the recursion k times. As $v_k \to 0$ exponentially fast, $\exists \varrho \in (0,1)$ s.t. for sufficiently large k,

$$\varrho^k \geq \prod_{j=1}^{k} \frac{(1-\varepsilon_j)|\det \mathbf{J}^{\mathrm{u}}|}{\mu_{j-1}}, \tag{20}$$

$$\begin{aligned} \Rightarrow \quad & \frac{1}{k} \sum_{j=1}^{k} \log_2 \mu_{j-1} \geq \log_2 |\det \mathbf{J}^{\mathrm{u}}| \\ & + \frac{1}{k} \sum_{j=1}^{k} \log_2 (1-\varepsilon_j) - \log_2 \varrho, \end{aligned}$$

$$\begin{aligned} \Rightarrow R \quad &\triangleq \liminf_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \log_2 \mu_{j-1} \geq \log_2 |\det \mathbf{J}^{\mathrm{u}}| \\ & + \liminf_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \log_2 (1-\varepsilon_j) - \log_2 \varrho, \quad (21) \\ &= \log_2 |\det \mathbf{J}^{\mathrm{u}}| - \log_2 \varrho, \quad (22) \\ &> \log_2 |\det \mathbf{J}^{\mathrm{u}}| = \log_2 \left| \prod_{\eta \in \sigma(\mathbf{J}^{\mathrm{u}})} \eta \right|, \\ &= \sum_{\eta \in \sigma(\mathbf{A}) : \, |\eta| \geq 1} \log_2 |\eta|, \quad (23) \end{aligned}$$

where (22) follows since $\varepsilon_j \to 0$ and the inequality in (23) from the fact that $0 < \varrho < 1$ strictly. This completes the proof of necessity.

Note that if the definition of stability used was weakened to uniform *asymptotic* stability, then the argument still applies but the strictness of the inequality in (23) is lost, since ϱ may not be strictly less than 1.

4 Tightness of Bound

The final step in proving Theorem 1 is to establish that the bound (6) is achievable, i.e. there exist coding and control schemes with asymptotic average data rates arbitrarily close to it that still achieve uniform exponential stability. In order to do so a specific coder-controller will be constructed and analysed.

The basis of the scheme is a two-phase strategy. In the first phase, a large but finite data rate is used to force the system into a specified b_0-neighbourhood of the origin in finite time. In the second phase, a more refined coder-controller is used to drive it exponentially fast towards the origin. This involves allocating each locally unstable component of $\mathbf{z}_k$ a an effective data rate roughly proportional to the log-magnitude of its governing local eigenvalue, while ignoring locally stable modes. Clearly, the asymptotic average data rate will be determined only by the second phase, since the first phase is completed in finite time with a constant data rate. It will then be shown that, by choosing a sufficiently small b_0, the asymptotic average data rate of the second phase can be made arbitrarily close to the RHS of (6).

Note that this scheme is not proposed as a practical control law, as issues such as performance, robustness and complexity would then need to be considered. It is intended only to demonstrate that the data rate lower bound (6) can be approached arbitrarily closely from above, making it the infimum exponentially stabilising data rate.

4.1 Phase One

By the controllability assumption A3, $\forall \varepsilon, l > 0$, $\exists N \in \mathbb{Z}_+, U > 0$ such that $\forall \|\mathbf{z}_0 - \mathbf{z}_*\| \leq l$, there is a control sequence $\tilde{\mathbf{v}}_{N-1}$ of length N and uniformly bounded by U that takes the system state *without communication constraints* to within ε of the fixed point, $\|\mathbf{z}_N^{\mathrm{nominal}} - \mathbf{z}_*\| \leq \varepsilon$. In principle, these controls are determined solely by the initial state, which is fully observed by the coder and can then pre-calculate them. One strategy the coder can then use is to

1. overbound the m-dimensional ball of radius U and centre $\mathbf{u}_*$ by a cube centred at $\mathbf{u}_*$ with sides of length $2U$ and partition this into μ_0 indexed, identical sub-cubes.

2. at time k, transmit the index s_k of the subcube which contains the nominal control signal $\mathbf{v}_{k+1}$,

Upon receiving this index, the controller then applies $\mathbf{u}_k =$ the midpoint of this sub-cube. As $\|\mathbf{u}_k - \mathbf{v}_k\| \leq U/\mu_0^{1/m}$, the continuity of $\mathbf{f}$ w.r.t. control implies that that $\forall \theta > 0, \exists \mu_0$ sufficiently large such that $\|\mathbf{z}_N - \mathbf{z}_N^{\mathrm{n}}\| \leq \theta$; hence for any permitted initial state,

$$\|\mathbf{z}_N - \mathbf{z}_*\| \leq \|\mathbf{z}_N - \mathbf{z}_N^{\mathrm{nominal}}\| + \|\mathbf{z}_N^{\mathrm{nominal}} - \mathbf{z}_*\| \leq \theta + \varepsilon$$

I.e. the state can be brought arbitrarily close to the origin in finite time N by using a constant, sufficiently large data rate $\log_2 \mu_0$.

4.2 Phase Two

Once the state is within some arbitrarily small neighbourhood of the origin, the coding and control is performed more carefully, by taking explicit account of the local dynamics. First, recall that the real Jordan form $\mathbf{J}$ of the Jacobian of $\mathbf{f}$ w.r.t. state at the set-point has a block diagonal structure

$$\mathbf{J} \equiv \mathrm{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_r) \in \mathbb{R}^{n \times n}.$$

Each block $\mathbf{J}_i \in \mathbb{R}^{d_i \times d_i}$ has either one distinct, real eigenvalue η_i of multiplicity $m_i = d_i$, or two distinct, complex conjugate eigenvalues $\eta_i, \overline{\eta}_i$ of multiplicities $m_i = d_i/2$. The components of the transformed state vector $\mathbf{z}_k$ corresponding to the block $\mathbf{J}_i$ are denoted $\mathbf{z}_k^{(i)} \in \mathbb{R}^{d_i}$. A further property of these blocks that will be used later is that $\exists\kappa > 0$ s.t.

$$\|\mathbf{J}_i^\tau\| \leq \kappa\tau^{d_i-1}|\eta_i|^\tau, \quad \forall N \in \mathbb{N}. \qquad (24)$$

This states that powers of a Jordan block grow exponentially according to the magnitude of its eigenvalue, with possibly an extra polynomial factor arising from multiplicity; see [7], pg. 138.

The coder to be used in the second phase can now be constructed. For convenience, the time index k is reset to zero, so that $\|\mathbf{z}_0 - \mathbf{z}_*\| \leq b_0$ a number that can be made arbitrarily small by increasing the data rate in the previous phase. Let R_0 be any number that satisfies (6) and divide times $k \in \mathbb{Z}_+$ into *epochs* $[j\tau, \ldots, (j+1)\tau - 1]$, $j \in \mathbb{Z}_+$, of some uniform integer duration τ. At time $k = j\tau$, suppose b_j is a pre-defined, uniform bound such that $\|\mathbf{z}_{j\tau} - \mathbf{z}_*\| \leq b_j, \forall \|\mathbf{z}_0 - \mathbf{z}_*\| \leq b_0$. The way in which $\{b_j\}_{j\in\mathbb{Z}_+}$ is generated will be specified later. Overbound this region by an n-dimensional cube centred at $\mathbf{z}_*$ with sides of length $2b_j$. Then partition this cube by dividing each coordinate axis corresponding to a component of $\mathbf{z}_{j\tau}^{(i)}$ into M_i intervals of equal length,

$$M_i \triangleq \lfloor|\xi\eta_i|^\tau\rfloor + 1 \;\text{ for } |\eta_i| \geq 1, M_i \triangleq 1 \;\text{ for } |\eta_i| < 1, \quad (25)$$

where $\lfloor\cdot\rfloor$ denotes rounding down and the parameter $\xi > 1$ is selected to satisfy

$$\begin{aligned} 0 < d\log_2\xi < R_0 - \sum_{|\eta_i|\geq 1} d_i \log_2|\eta_i| \\ \equiv R_0 - \sum_{|\eta|\in\sigma(\mathbf{A}):|\eta|\geq 1} \log_2|\eta|. \qquad (26)\end{aligned}$$

Note that as the right-hand side (RHS) is guaranteed positive, candidates for ξ always exist. The total number of subcuboids thus formed is $\prod_{i=1}^r M_i^{d_i}$, so index them in a predefined way by the integers $0, \ldots, \mu_{j\tau} - 1$, where $\mu_{j\tau} \triangleq \prod_{1\leq i\leq r} M_i^{d_i}$. At time $j\tau$, transmit the index $s_{j\tau}$ of the one which contains $\mathbf{z}_{j\tau}$. At remaining times in the epoch $j\tau + 1 \leq k \leq (j+1)\tau - 1$, set $\mu_k = 1$, i.e. transmit no information. Clearly, the asymptotic average data rate over both phases is determined only by the average data rate over this second phase, so

$$\begin{aligned} R &= \frac{1}{\tau}\sum_{|\eta_i|\geq 1} d_i\log_2\left(\lfloor|\xi\eta_i|^\tau\rfloor + 1\right) \leq \frac{1}{\tau}\sum_{|\eta_i|\geq 1} d_i\log_2\left(2|\xi\eta_i|^\tau\right), \\ &= \frac{d}{\tau} + d\log_2\xi + \sum_{|\eta_i|\geq 1} d_i\log_2|\eta_i| < R_0 \end{aligned}$$

by (26), for sufficiently large τ.

Now consider the controller at the other end of the channel, which receives the symbol $s_{j\tau}$ at time $j\tau + 1$. As it also knows the uniform bound b_j and the number of intervals M_i, $i = 1, \ldots, r$, it then knows which subcuboid $\mathbf{z}_{\tau j}$ lies in and uses its centre as an estimate $\mathbf{q}_j$. Hence

$$\|\mathbf{z}_{j\tau}^{(i)} - \mathbf{q}_j^{(i)}\| \leq \left[\sum_{h=1}^{d_i}\left(\frac{b_j}{M_i}\right)^2\right]^{1/2} = \frac{\sqrt{d_i}}{M_i}b_j. \qquad (27)$$

It then calculates the next n control signals $\mathbf{u}_{\tau j+1}, \ldots, \mathbf{u}_{\tau j+n}$ by using the controllability of $(\mathbf{A}, \mathbf{B})$, and hence of $(\mathbf{J}, \mathbf{TB})$, to force the linearised system with nominal initial state $\mathbf{q}_j$ to the origin in $n+1$ steps, i.e. by solving

$$\begin{aligned} &\sum_{k=j\tau+1}^{j\tau+n} \mathbf{J}^{\tau j+n-k}\mathbf{TB}(\mathbf{u}_k - \mathbf{u}_*) \\ &\equiv \mathbf{W}(\mathbf{y}_j - \mathbf{y}_*) = -\mathbf{J}^n\mathbf{T}(\mathbf{q}_j - \mathbf{z}_*), \;\; \forall j \in \mathbb{Z}_+, \quad (28)\end{aligned}$$

where $\mathbf{W} \triangleq [\mathbf{TB} \;\; \mathbf{JTB} \cdots \mathbf{J}^{n-1}\mathbf{TB}] \in \mathbb{R}^{n\times nm}$ and $\mathbf{y}_j \triangleq [\mathbf{u}'_{j\tau+n} \;\; \mathbf{u}'_{j\tau+n-2} \cdots \mathbf{u}'_{j\tau+1}]' \in \mathbb{R}^{nm}$. The remaining control signals in the epoch are set to $\mathbf{u}_*$

Note that as the controllability matrix $\mathbf{W}$ has rank n, it possesses n linearly independent columns $\in \mathbb{R}^n$ and only the corresponding n scalar components of the stacked control vector $\mathbf{y}_j$ are needed. If the inverse of the matrix formed by these columns is padded with $nm - n$ null rows, corresponding to the unnecessary components of $\mathbf{y}_j$, to form $\mathbf{V} \in \mathbb{R}^{nm\times n}$, then the stacked control may be expressed more explicitly as

$$\mathbf{y}_j - \mathbf{y}_* = -\mathbf{V}\mathbf{J}^n(\mathbf{q}_j - \mathbf{z}_*), \qquad (29)$$

i.e. a linear function of $\mathbf{q}_j - \mathbf{z}_*$.

The crucial remaining step is to determine how to update the uniform upper bound b_j from one epoch to the next. In the following a recursion for b_j will be sought which decays exponentially to zero for a sufficiently large but fixed epoch duration τ. As $b_j \geq \|\mathbf{z}_{j\tau} - \mathbf{z}_*\|$ by definition, this will effectively complete the proof of Theorem 1.

First observe that, as $\mathbf{q}_j \in \mathbb{R}^n$ lies in a cube of sides $2b_j$ centered at $\mathbf{z}_*$, $\|\mathbf{q}_j - \mathbf{z}_*\| \leq \sqrt{n}b_j$. In addition, by (29) $\exists C \triangleq \|\mathbf{V}\mathbf{J}^n\|$ independent of τ and b_j s.t. $\|\mathbf{u}_k - \mathbf{u}_*\| \leq Cb_j$ for all times k in the jth epoch. Now consider the map $\mathbf{f}$ iterated τ times from some initial state $\mathbf{z}$ and with inputs $\mathbf{v}_0, \ldots, \mathbf{v}_{\tau-1}$, denoted $\mathbf{f}_{\mathbf{v}_{\tau-1}}\cdots\mathbf{f}_{\mathbf{v}_0}(\mathbf{z})$ for convenience. By the continuous differentiability of $\mathbf{f}$, it follows that $\forall \|\mathbf{z} - \mathbf{z}_*\| \leq b$, $\|\mathbf{v}_t - \mathbf{u}_*\| \leq Cb$, $t = 0, \ldots, \tau - 1$, $\exists\zeta(b)$ s.t.

$$\begin{aligned} b\zeta(b) \;\; \geq \;\; &\Big\|\mathbf{f}_{\mathbf{v}_{\tau-1}}\cdots\mathbf{f}_{\mathbf{v}_0}(\mathbf{z}) - \mathbf{z}_* - \mathbf{J}^\tau(\mathbf{z} - \mathbf{z}_*) \\ &- \sum_{t=0}^{\tau-1}\mathbf{J}_i^{\tau-1-t}\mathbf{TB}(\mathbf{v}_t - \mathbf{u}_*)\Big\| \end{aligned} \qquad (30)$$

where $\zeta(b)$ may depend on τ but $\to 0$ with b. Substituting $\mathbf{z} = \mathbf{z}_{j\tau}$, $\mathbf{v}_t = \mathbf{u}_{j\tau+t}$, rearranging and looking at each ith local mode,

$$\begin{aligned} \|\mathbf{z}_{(j+1)\tau}^{(i)} - \mathbf{z}_*^{(i)}\| &\leq \|\mathbf{J}_i^\tau\|\|\mathbf{z}_{j\tau}^{(i)} - \mathbf{q}_j^{(i)}\| + \zeta(b_j)b_j, \\ &\leq \kappa\tau^{d_i-1}|\eta_i|^\tau\frac{\sqrt{d_i}b_j}{M_i} + \zeta(b_j)b_j. \quad (31)\end{aligned}$$

$$\begin{aligned} &\Rightarrow \|\mathbf{z}_{(j+1)\tau} - \mathbf{z}_*\| \\ &\leq \sqrt{r}\kappa\max_{1\leq i\leq r}\left\{\tau^{d_i-1}|\eta_i|^\tau\frac{\sqrt{d_i}}{M_i}\right\}b_j + \sqrt{r}\zeta(b_j)b_j, \quad (32) \\ &\equiv [\beta(\tau) + \sqrt{r}\zeta(b_j)]b_j =: b_{j+1}, \quad (33)\end{aligned}$$

where (31) follows from (24) and (27), and (32) from the triangle inequality.

Now, consider the b_j-independent term $\beta(\tau)$ on the RHS of (33). If $|\eta_i| < 1$, $M_i = 1$ and $\tau^{d_i-1}|\eta_i|^\tau \to 0$ as $\tau \to \infty$. If $|\eta_i| \geq 1$, $M_i \geq |\xi\eta_i|^\tau$ from (25), so

$$\tau^{d_i-1}|\eta_i|^\tau \frac{\sqrt{d_i}}{M_i} \leq \sqrt{d_i}\tau^{d_i-1}\xi^{-\tau} \to 0 \quad \text{as } \tau \to \infty.$$

Hence, $\beta(\tau)$ can be made arbitrarily small, independently of b_j. Select some value for τ large enough that $\beta(\tau) < 1$. As for any fixed τ, $\zeta(b) \to 0$, and further recalling that b_0 can be made arbitrarily small by the first phase of the coding and control scheme, set b_0 so that $\forall b \leq b_0$, $\beta(\tau) + \sqrt{r}\zeta(b) \leq$ some selected $\chi < 1$. It then follows that $b_1 \leq \chi b_0 < b_0$, and by induction it can be established that $b_{j+1} \leq \chi b_j < b_j$.

Hence $b_j \to 0$ exponentially fast and, recalling that $\|\mathbf{u}_{j\tau} - \mathbf{u}_*\| \leq Cb_j$, the closed loop system is then uniformly exponentially stable in state and control at times $k = j\tau$. The continuity of $\mathbf{f}$ and the linear dependence of $\mathbf{u}_k - \mathbf{u}_*$ on $\mathbf{q}_j - \mathbf{z}_*$ can then be used to show that it is uniformly exponentially stable as $k \to \infty$ over the integers.

5 Conclusion

In this paper the problem of uniformly exponentially stabilising a deterministic, nonlinear dynamical system under a feedback data rate constraint was formulated and investigated. By using volume-partitioning arguments and local Jordan forms, the infimum asymptotic average data rate for the system to be stabilisable was derived, in terms of the unstable eigenvalues of the dynamic map at the desired set-point. Connections to the notion of topological feedback entropy have been explored elsewhere [12].

References

[1] J. Baillieul. Feedback designs in information-based control. In *Stoch. The. Contr. Proc. Worksh.*, pages 35–57, Uni. Kansas, USA, Oct 2001. Springer.

[2] R. W. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Trans. Autom. Contr.*, 45(7):1279–89, 2000.

[3] D. F. Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE Trans. Autom. Contr.*, 35:916–24, 1990.

[4] N. Elia and S. K. Mitter. Stabilization of linear systems with limited information. *IEEE Trans. Autom. Contr.*, 46:1384–400, 2001.

[5] F. Fagnani and S. Zampieri. Stabilizing quantized feedback with minimal information flow: the scalar case. In *Proc. 15th Int. Symp. Math. The. Netw. Sys.*, U. Notre Dame, USA, Aug 2002.

[6] J. Hespanha, A. Ortega, and L. Vasudevan. Towards the control of linear systems with minimum bit-rate. In *Proc. 15th Int. Symp. Math. The. Netw. Sys.*, U. Notre Dame, USA, Aug 2002.

[7] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[8] H. Ishii and B. A. Francis. Quadratic stabilization of sampled-data systems with quantization. In *Proc. 15th IFAC World Cong.*, Barcelona, 2002.

[9] G. N. Nair and R. J. Evans. Stabilization with data-rate-limited feedback: tightest attainable bounds. *Sys. Contr. Lett.*, 41(1):49–56, Sep 2000. Also available at www.sciencedirect.com.

[10] G. N. Nair and R. J. Evans. Mean square stabilisability of stochastic linear systems with data rate constraints. In *Proc. 41st IEEE Conf. Dec. Contr.*, pages 1632–7, Las Vegas, USA, Dec 2002.

[11] G. N. Nair and R. J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39:585–93, 2003. Also available at www.sciencedirect.com. Short version published in Proc. 15th IFAC World Congress, July 2002.

[12] G. N. Nair, R. J. Evans, I. Mareels, and W. Moran. Topological feedback entropy and nonlinear stabilization. Under review, 2003.

[13] I. R. Petersen and A. V. Savkin. Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel. In *Proc. 40th IEEE Conf. Dec. Contr.*, pages 304–9, 2001.

[14] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. Jour.*, 1948. Reprinted in 'Claude Elwood Shannon Collected Papers', IEEE Press, 1993.

[15] S. Tatikonda and S. Mitter. Control under communication constraints. In *Proc. 38th Ann. Allerton Conf. Comm. Contr. Comp.*, pages 182–90, Oct 2000.

[16] T. Ward. *Entropy of Compact Group Automorphisms*. Dept. Mathematics, Uni. East Anglia, UK, 1994. Based on notes for course at Ohio State Uni., USA. Available at $http://www.mth.uea.ac.uk/\ h720/lecture\_notes/$.

[17] W. S. Wong and R. W. Brockett. Systems with finite communication bandwidth constraints II: stabilization with limited information feedback. *IEEE Trans. Autom. Contr.*, 44:1049–53, 1999.